METHOD AND APPARATUS FOR ROBUST EFFICIENT PARSING

ABSTRACT OF THE DISCLOSURE

The present invention provides a method for

5   improving the efficiency of parsing text. Aspects of the invention include representing parse tokens as integers where a portion of the integer indicates the location in which a definition for the token can be found. In a further aspect, an integer representing a

10   token points to an array of tokens that can be activated by the token. In another aspect, a list of pointers to partial parses is created before attempting to parse a next word in the text string. The list of pointers includes pointers to partial

15   parses that are expecting particular semantic tokens. A fourth aspect of the invention utilizes a data structure to list the semantic tokens that have been fully parsed for each span in the input text segment. When a token is fully parsed, the list is accessed to

20   determine if the new token should be discarded.